



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 4, April 2022

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.118



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com

Car Price Prediction using Linear Regression Technique of Machine Learning

V.C.Sanap¹, Mohammed Munawwar Rangila², Sufiyaan Rahi³, Samiksha Badgujar⁴, Yashodhan Gupta⁵

Lecturer, Dept. of Computer Engineering, A.I.S.S.M.S. Polytechnic Pune, Maharashtra, India¹

Student, Dept. of Computer Engineering, A.I.S.S.M.S. Polytechnic Pune, Maharashtra, India^{2, 3, 4, 5}

ABSTRACT: In this fast-moving generation, the present study proposes the newer concept of predicting the prices of certain items. With an idea and motivation to help everyone we came up with a solution to get an appropriate estimate of one's car using Machine Learning Techniques which will save a lot of time and money. A car price prediction has been a high interest research area, as it requires noticeable effort and knowledge of the field expert. Considerable number of distinct attributes are examined for the reliable and accurate prediction. To build a model for predicting the price of used cars in, we applied one of the machine learning techniques i.e., Linear Regression. Using linear regression, there are multiple independent variables but one and only one dependent variable whose actual and predicted values are compared to find precision of results. Our paper proposes a system where price is dependent variable which is predicted, and this price is derived from factors like distance driven, car purchase year etc.

KEYWORDS: Car Price Prediction, Linear Regression, Data Cleaning, Data Understanding.

I. INTRODUCTION

In our day to day lives everyone buys and sells a car every day. Now there are limited facilities and applications to get an appropriate price for one's car. Now we need an application to get an estimate value of the car. We use a linear regression algorithm for this purpose. This object comprises of Car as an object and the price for every car entered in the dataset. The dataset we work on in this venture is imaginary i.e., not real. Linear Regression allows us to compare our entered data to the existing data in the dataset and gives us an estimate value. Car works as a main object. The attributes such as Company Name, Kilometres Driven, Fuel Type and Year of Purchase etc are upon which the comparison is performed using Machine Learning Techniques. Also, further on discussing the imperatives, we have utilized composite configuration design, which make it well extendable to include or uproot as numerous.

II. RELATED WORK

Researchers more often predict prices of products using some previous data and so did Pudaruth [1] who predicted prices of cars in Mauritius and these cars were new rather second hand. He used multiple linear regression, k-nearest neighbors, naïve Bayes and decision trees algorithm in order to predict the prices. The comparison of prediction results from these techniques showed that the prices from these methods are closely comparable. However, it was found that decision tree algorithm and naïve bayes method were unable to classify and predict numeric values. Pudaruth's research also concluded that limited number of instances in data set do not offer high prediction accuracies [1]. Multivariate regression model helps in classifying and predicting values of numeric format. Kuiper [2] used this model to predict price of 2005 General Motor (GM) cars. The price prediction of cars does not require any special knowledge so the data available online is enough to predict prices like the data available on www.pakwheels.com. Kuiper [2] did the same i.e. car price prediction and introduced variable selection techniques which helped in finding which variables are more relevant for inclusion in model. He encouraged students to use different models and find how checking model assumptions work. Another similar research by Listiani [3] uses Support Vector Machines (SVM) to predict the prices of leased cars. This research showed that SVM is far more accurate in predicting prices as compared to the multiple linear regression when a very large dataset is available. SVM also handles high dimensional data better and avoids both the under-fitting and over-fitting issues. Genetic algorithm is used by Listiani [3] to find important features for SVM. However, the technique does not show in terms of variance and mean standard deviation why SVM is better than simple multiple regression.

III. METHODOLOGY

Our research aims to develop a good machine learning model which offers accurate prediction of a car's price using linear regression. In order to do this, we needed some previous data set of pre-owned cars for estimating the car price on the basis of some attributes like price, year of purchase, distance travelled, model etc. These attributes are taken as independent variables whereas the car price is dependent on these variables.

The equation of linear regression which combines our attributes and displays the result is based on:

$$Y = \beta_0 + \beta_1 X$$

In the above represented equation β_0 , β_1 represents the regression coefficients, Y as the output or required variable and X shows the input. Relation in case of single input is represented by the above equation. For multiple inputs by the user the given below equation is used.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

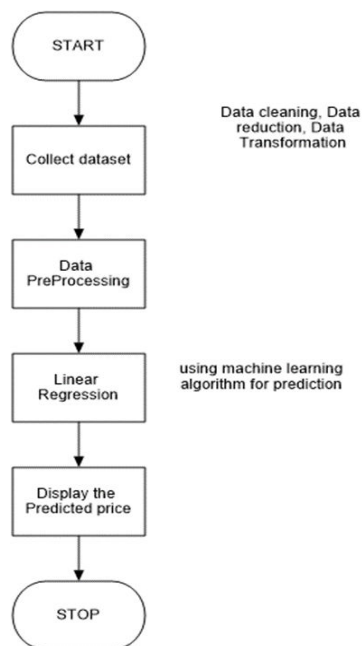
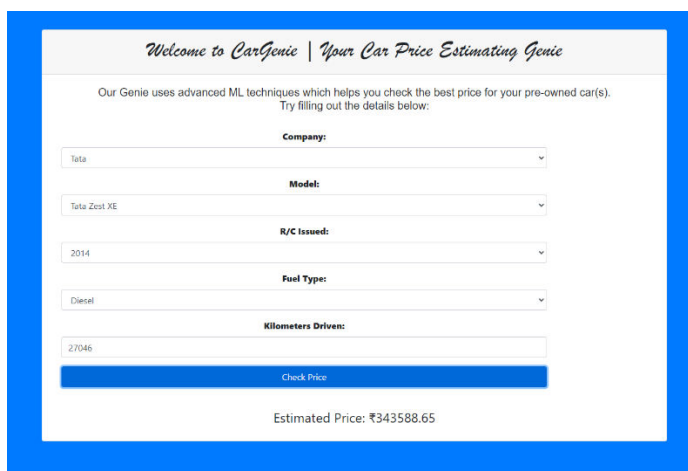
There are many attributes and features that impact the price of a car which ultimately generates large sets of data which leads to complexity in analysing it. The focus of our research is to build such a model which is capable of dealing with high complexity algorithms and gives accurate results irrespective of the magnitude of the data set.

The input data set which is collected from online website Quikr in the time span of three to four months ranging from top shelf cars to daily commute cars. The collected data included variable inputs like colour, rims, body type, transmission, engine capacity, register city, power steering, alloys, mileage etc. which is super un-cleaned data

Once the data collection was over we cleaned the data which gave us the attributes based on which we have developed our system upon. We processed the data using multiple linear regression technique for predicting the price of a particular car.

IV. EXPERIMENTAL RESULTS

The proposed model of our Car Price Prediction is shown below. The model involves attributes like the car's company name, model of the car, rc issued which is the year of purchase, fuel type and kilometres driven. First as shown in the image (a) The user selects the company of the car he owns. The company names are fetched from the dataset we had provided in the backend for the user to successfully select the company name. (b) The user then selects the car model of the particular company. The car model is also fetched from the dataset. (c) R/C Issued is the year in which the car was purchased. (d) The fuel type of the car is to be selected. (e) Kilometers driven Lets the user input how much the car has travelled. (f) Finally upon clicking on the check price button after successfully filling the attribute fields the estimated car price is displayed which gives the user an idea at which the user can sell the car or lets him determine the value of the car.

Welcome to CarGenie | Your Car Price Estimating Genie

Our Genie uses advanced ML techniques which helps you check the best price for your pre-owned car(s). Try filling out the details below:

Company:

Model:

R/C Issued:

Fuel Type:

Kilometers Driven:

Estimated Price: ₹343588.65



V. CONCLUSION

Our Car Price Prediction System is used to get an accurate estimation of a car we would want to get the resale value of. Our Car Price Prediction System takes five attributes as inputs and based on those inputs we get an estimated price. We use Machine Learning Techniques such as Linear Regression to get our predicted price. Our Car Price Prediction System is reliable and efficient

REFERENCES

1. GONGGI, S., 2011. New model for residual value prediction of used cars based on BP neural network and non-linear curve fit. In: Proceedings of the 3rd IEEE International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), Vol 2. pp. 682-685.
2. WU, J. D., HSU, C. C. AND CHEN, H. C., 2009. An expert system of price forecasting for used cars using adaptive neuro-fuzzy inference. Expert Systems with Applications. Vol. 36, Issue 4, pp. 7809-7817
3. LISTIANI, M., 2009. Support Vector Regression Analysis for Price Prediction in a Car Leasing Application. Thesis (MSc). Hamburg University of Technology-deep-learning
4. Richardson, M. S. (2009). Determinants of used car resale value. Retrieved from: <https://digitalcc.coloradocollege.edu/islandora/object/coccc%3A1346> [accessed: August 1, 2018.]
5. Noor, K., & Jan, S. (2017). Vehicle Price Prediction System using Machine Learning Techniques. International Journal of Computer Applications, 167(9), 27-31.
6. Pudaruth, S. 2014. "Predicting the Price of Used Cars Using Machine Learning Techniques", International Journal of information & Computation Technology, 4(7), p.753-764.
7. Minitab Express Support. Interpret all statistics for Predict.[Online] Available from: <http://support.minitab.com/en-us/minitab-express/1/help-and-how-to/modeling-statistics/regression/how-to/predict/interpret-the-results/all-statistics/>



INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.118



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details